DOCUMENT RESUME

ED 371 043 TM 021 747

AUTHOR Loup, Karen S.; And Others
TITLE Initial Development and Piloting of a Learning-Based, Classroom Assessment and Consultation System: New Perspectives on the Rhetoric of Improving Instruction in Higher Education Settings.
PUB DATE Apr 94
NOTE 47p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *College Instruction; Consultation Programs; *Continuation Students; Continuing Education; *Educational Assessment; Educational Improvement; Evaluation Methods; Feedback; Field Studies; Higher Education; Models; *Pilot Projects; Research and Development; *Test Construction; Test Reliability; Test Use; Test Validity
IDENTIFIERS *Student Assessment of Teaching and Learning

ABSTRACT

Results are reported of three years of research and development, piloting, and extended field testing of a classroom-based assessment and professional consultation system used to assess important teaching and learning variables in higher education contexts. Of particular interest is the focus of the total system on enhancing learning and newer assessment concerns such as teaching for development of higher-order thinking skills. In phase 1, the consultation/feedback model and the assessment instrument, the Student Assessment of Teaching and Learning (SATL), were developed. In phase 2 the SATL and the model were pilot tested with the inclusion of student and faculty perceptions of the instrument. Phase 3 studies of the validity and reliability of the SATL were conducted with 28 continuing-education classes. The majority of student participants were female (67.2%) most of whom were older than traditional college students. Data from 448 students were the basis for additional trials in phase 4 with nearly 1,000 students and 80 faculty members from a larger sample of evening-school classes. Results provide empirical support for the validity and reliability of the instrument and suggest that the instrument and the model provide valuable information on teaching and learning in higher education. Four tables present study data. Appendix A contains the course evaluation report profiles, and Appendix B contains the course evaluation report histograms. (Contains 32 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

BARBARA S. MOFFETT

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 371 043

Initial Development and Piloting of a Learning-Based, Classroom Assessment and Consultation System: New Perspectives on the Rhetoric of Improving Instruction in Higher Education Settings

Karen S. Loup
Rita R. Culross
Chad D. Ellett
Lynn Evans

Louisiana State University

Paper Presented at the Annual Meeting of the American Educational Research Association
New Orleans, Louisiana
April, 1994

BEST COPY AVAILABLE

TM021747

Initial Development and Piloting of a Learning-Based, Classroom Assessment and Consultation System: New Perspectives on the Rhetoric of Improving Instruction in Higher Education Settings

The quality of undergraduate teaching and learning environments on college campuses is a recent national concern. The recent call to "reaffirm teaching as the university's primary task" is echoed by the Carnegie Foundation for the Advancement of Teaching in its proposal to redefine "scholarship" in ways that emphasize the importance of improving undergraduate teaching and learning environments. Literature reviews by Purtle (1982) and Aleamoni (1987) reveal that considerable evidence supports the use of student ratings as one data source in the assessment of course characteristics and the quality of instruction, particularly when such information is used for formative evaluation purposes. While some learning environment instruments have been developed for use in higher education settings, (e.g., the College and University Classroom Environment Inventory (CUCEI), Fraser, Treagust, & Dennis, 1986), no measures are known that attempt to assess students' perceptions in higher education settings of the extent to which teaching and learning environment variables enhance learning, particularly n newer, important areas such as the development of higher order thinking skills. Fraser, Treagust, Williamson & Tobin (1987), citing the strong tradition of classroom environment research using student perceptions measures at the primary and secondary school levels, suggest that this research should form a basis for the development of classroom learning environment perceptions measures suitable for use in higher education contexts. Marsh & Dunkin (1992) and Menges (1992) also recommend that researchers in teaching in higher education contexts look to the more advanced research base of effective teaching in elementary and secondary schools as a starting point in investigating teaching and learning in higher education classes.

During the last ten to fifteen years, much has been learned about important teaching behaviors and their contribution to student engagement and the enhancement of student learning (Brophy, 1986; Porter & Brophy, 1988) and teaching and learning scholars are moving forward with additional models of effective teaching based upon explications and understandings of teacher knowledge (Shulman, 1986; 1987) and teacher expertise (Berliner, 1986; 1989). Also, recent research on the study of learning environments has revealed that student outcomes in various educational settings depend to a large extent on the nature of the actual classroom environment (Fraser & Fisher, 1983; Fraser et al., 1987; Marcelo, 1990). Logan and Ellett (1988) suggest that even though higher education settings are somewhat different teaching and learning contexts than public school classrooms, it seems reasonable that the recent process product-research literature, literature on classroom learning environments and research on student learning, should provide an initial focus for the development of an instrument to measure effective teaching and learning environments in higher education contexts. In addition, the extent to which students' learning is enhanced in higher education classroom contexts is an important assessment concern, from both formative and summative evaluation perspectives. This call for more comprehensive views of assessment in higher education settings is at odds with the more traditional and more narrow use of student ratings of courses alone.

Recent meta analyses (Menges & Brinko, 1986) and experimental studies of the effects of systematic feedback also note (Marsh & Roche, 1993) reveal that feedback from student ratings alone produces small improvements in teaching. However, when this feedback is accompanied by professional consultation, the effects on teaching improvement are greatly enhanced. Menges and Brinko (1986) also note that among studies that include consultation, there is considerable variation in effect sizes associated with systematic consultation procedures. Brinko (1991) compares five models of interaction which are applicable to the consultation

process. Each model posits a different role for the consultant as well as for the participant. For example, a consultant's role might be one of an expert, a facilitator of change, a diagnostician, a psychological counselor, or a challenger. Participants may assume roles such as a seeker of personal growth, a receiver of information, a content expert, and the like in interactions during the consulting process. Thus, selecting an appropriate consultation/feedback model to accommodate individual differences among faculty in higher education settings is an important concern to enhance teaching and learning environments for students.

PURPOSE

The purpose of this paper is to report the results of three years of research and development, piloting, and extended field testing of a classroom-based assessment and professional consultation system used to assess important teaching and learning variables in higher education contexts. Of particular interest is the focus on the total system on enhancing learning and newer assessment concerns (e.g., teaching for development of higher order thinking skills). The assessment consultation system was developed and piloted as: 1) an alternative to evaluation instruments traditionally used in higher education classrooms that are designed to measure more narrowly focused instructor or course evaluation characteristics; 2) a system reasonably grounded in the emerging research literatures reflecting linkages between elements of effective teaching and student learning; and 3) an assessment framework to tap student perceptions in a higher education context of key elements reflected in a classroom-based, on-the-job observation system originally designed to evaluate teaching and learning in public school classrooms in Louisiana (Ellett, 1990; Ellett, Loup, & Chauvin, 1991a; 1991b).

METHODOLOGY

This study was completed in three phases of R&D activity as follows: Phase I: Assessment instrument and consultation/feedback model development; Phase II; Initial piloting

of the system to examine utility/feasibility for use in undergraduate classrooms; Phase III: Studies of validity and reliability of instrumentation; Phase IV: Piloting of a comprehensive, collaborative/feedback consultation model with higher education faculty. Each of these phases is described in the following sections.

Phase I: Instrument and Model Development

Assessment, Consultation, and Feedback Model

A comprehensive, classroom-based assessment, consultation and feedback model was designed that was grounded in the literatures on teaching and learning cited above (Evans & Ellett, 1990) and a collaborative process and consultation model (Brinko, 1991). The model was enhanced by the use of a combination of three triangulated data sources to obtain classroom-based assessment information to be used for formative evaluation purposes: 1) student perceptions; 2) observations of external observers; and 3) faculty self-assessments. Assessment data from these three sources are used to identify targets for consultation and discussion in formative evaluation/professional development conferences with faculty.

Instrumentation

Initially, some 30 different instruments targeting student evaluations of instruction and courses in higher education classes were reviewed as to content and response format. None of these instruments was found to have a primary focus on the actual processes and interactions of teaching and learning within the higher education classroom learning environment context. With few exceptions, these instruments focused upon the evaluation of course content, activities, materials and/or personal characteristics of the instructor. Moreover, few of the instruments reviewed appeared to be useful for the purposes of providing feedback for formative evaluation

and instructional improvement. Based upon this review and the need for collecting assessment data to be used for formative evaluation purposes, a student perceptions assessment instrument for use in higher education learning environments was developed for this study. Instrument content and assessment formats were derived from the following sources: 1) a synthesis of higher education course evaluation and student perceptions measures literatures; 2) existing models and perceptions measures in the study of classroom learning environments (e.g., Moos, 1974; Fraser, Anderson & Walberg, 1987; Fraser & Fisher, 1983; Fraser et al., 1987, Ellett, Loup & Chauvin, 1991; Loup, Ellett & Chauvin, 1992); 3) a syntheses of process-product research findings (Brophy, 1986; Porter & Brophy, 1988); 4) a review and synthesis of large-scale teacher performance assessment instruments developed to evaluate teaching in public school contexts in Georgia, Florida, South Carolina, North Carolina, Tennessee, and Texas (Ellett, Garland & Logan, 1987); and 5) more recent assessment concerns such as the teaching of thinking skills and learning equity reflected in the System for Teaching and learning Assessment and Review (STAR) (Ellett, Loup & Chauvin, 1991b). The resulting 68-item draft of a student perceptions measure termed the Student Assessment of Teaching and Learning (SATL)(Evans & Ellett, 1992) was designed to assess key elements of the interactive nature of classroom teaching and learning environments in higher education settings.

The initial draft of the 68 assessment items of the SATL consisted of indicators of effective teaching and learning grouped by four instrument categories:

I. Preparation and Classroom Management (11 indicators)

II. Interpersonal Skills (10 indicators)

III. Enhancement of Learning (24 indicators)

IV. Student Evaluation Practices (10 indicators)

Assessment indicators defining various Teaching and Learning Components of the STAR (Ellett, et. al, 1991b) were used to develop the initial pool of assessment indicators for the SATL for the first three categories. The STAR assessment indicators were slightly reworded where appropriate to make them more applicable to assessments in higher education contexts. For example, the original STAR indicator statement...Enthusiasm about teaching, learning and the subject taught is communicated throughout the lesson...was rewritten in the SATL as a set of two assessment indicator statements...The professor is enthusiastic about teaching this course...and...The professor is enthusiastic about students' learning in this course. The SATL assessment indicators collectively reflect a complex set of classroom activities, conditions, events and/or behaviors that define core elements of teaching and learning rather than specific prescriptive teaching behaviors or targeted course characteristics. A list of the SATL assessment indicators classified by each of four Assessment Domains can be found in Table 1.0

The response format for each SATL assessment indicator is a three-point, forced choice scale ranging from: 1 = INEFFECTIVE (does not positively enhance student learning); 2 = EFFECTIVE (for the most part, positively enhances student learning; and 3 = HIGHLY EFFECTIVE (consistently enhances student learning to a high degree). The response format represents a considerable change from typical course evaluation rating scales (e.g., Never, Sometimes, Always...or... Needs Improvement, Satisfactory, Excellent). In assessing each SATL indicator, students are asked to reflect on the extent to which the indicator as evidenced in the class/course contributed to the enhancement of their individual learning. Thus, the student response set for assessing each of the SATL indicators is learning-centered and outcomes-based.

In addition, for purposes of the pilot study, students were also asked to respond to ten items that reflected the degree to which each of ten types of learning (e.g., critical analysis and/or problem solving, creative thinking, learning factual information, etc.) was emphasized in the particular course being evaluated. In a third part of the SATL, students were also asked to provide global, summary evaluation judgements of: 1) the quality of teaching; 2) the contribution of the particular course to their learning; and 3) a predicted final course grade.

Phase II: Initial Piloting

The SATL and the consultation model was initially piloted to develop six case studies of teaching and learning in classes taught by undergraduate teaching assistants (Evans & Ellett, 1990). The pilot of the assessment model included collection of student and faculty self-perceptions of teaching and learning using the SATL. Systematic classroom observations were completed by external observers who had been trained in the used of the System for Teaching and learning Assessment and Review (STAR). Syntheses of SATL and STAR data were used as a basis for ensuing and ongoing consultations with the teaching assistants to improve teaching and learning in their classes.

Phase III: Studies of the Validity and Reliability of the SATL

Sample

During the spring of 1992, the SATL was administered to 34 Evening School/Continuing Education, off-campus courses (classes) offered during the spring semester, 1992 through Louisiana State University (LSU). These classes were located in 12 different education centers distributed throughout Louisiana. A wide range of subject areas was represented in these courses, including classes drawn from the biological and social sciences, math, English, foreign languages,

education and foundations classes (e.g., philosophy). The academic level of courses ranged from freshman year courses through doctoral-level special topics courses. Class sizes varied from nine to 55 students.

Data Collection Procedures

The initial field trial of the SATL reported here was completed during the spring of 1992 with a random sample of students in Evening School classes on the LSU main campus using a cross-section of courses representing a wide variety of higher education contexts. One purpose of this initial field trial was to establish estimates of item and instrument category descriptive statistics and internal consistency reliabilities. A second purpose was to generate student perceptions data for examining construct validity characteristics of the SATL using a series of exploratory factor analyses. The SATL was confidentially administered to students in whole class groups and required approximately 15-20 minutes to complete.

Classes were randomly selected within various subject area categories (e.g., social sciences, biological sciences, humanities, English, etc.). Instructors/faculty in the sample were contacted by letter and asked to cooperate in requesting that their students complete the SATL. Completion of the SATL was described to students as an initial attempt to develop a more systematic procedure for evaluating students' perceptions of Evening School courses. It was explained that the focus of the study was on students' collective perspectives of Evening School courses and that the SATL data would be used to provide meaningful feedback to instructors for course improvement purposes. Subsequent class summaries of SATL data were considered to be confidential and these were not reported to various academic departments.

SATL data for the study were collected during the last two weeks of regularly scheduled

Evening School classes during the spring of 1992. Students were asked to complete the SATL instrument, a demographic sheet, and a comment/response form which solicited additional student perspectives about the instructor/course, the clarity of directions and items, appropriateness of items for the particular course, and any important aspects of teaching and learning that were not included in the SATL instrument, etc. Students were also asked to indicate the length of time required to complete all data collection tasks.

The SATL student perception instruments were distributed and returned by campus mail and U.S. mail. Follow-up calls were made to instructors after the end of the semester. Of the 34 classes selected in the original sample, 28 (82%) participated in the study. The majority of student participants were female (67.2%) and over the age of 25 (66.8%). Only 7.4% of students in the sample were of the traditional college age (18-21). Slightly more than one-half were enrolled in a degree program. Almost two-thirds were part-time students, and 47% had not taken a college course within one to three years of the current (spring, 1992) semester. For more than 22% of the students in the sample, the course evaluated was the first college course taken in 10 or more years.

Data Analyses

Data were available for various analyses for 448 students distributed across 28 different classes. A variety of data analyses were completed to examine the initial psychometric characteristics of the SATL. These included:

1. Descriptive statistical summaries for each item (assessment indicator)
2. A variety of principal component, oblique and orthogonal factor analyses
3. Alpha reliability analyses for each SATL Domain (Subscale)

4. Intercorrelations among SATL Domains (Subscales)

5. Intercorrelations among expected course grades and SATL Domains (Subscales)

Results

A variety of exploratory factor analyses were completed to examine the original content classification of SATL items into the four Assessment Domains. Results of a one-factor, principal components solution indicated that the 55 items comprising SATL Assessment Domains I, II and III meaningfully loaded on a single factor (range of loadings =.38 to .83), with 46 of 55 loadings exceeding .60. Table 1 summarizes the results of this analysis. Item/factor pattern loadings shown in the Table are correlation coefficients. This one-factor solution accounted for 45% of the total variation in the data.

Subsequently, a series of oblique and orthogonal solutions extracting from two to ten factors was completed. For oblique solutions, a large number of items cross-loaded on more than one factor. Results of a four-factor, orthogonal solution best depicted the alignment of SATL indicators with the various Assessment Domains. Factor loadings for this solution ranged from .36 to .77 with 21 loadings at or exceeding .60. Results of this analysis are shown in Table 2. This solution accounted for 56.23% of the total variance in the data, with 12 of 55 indicators (items) cross-loading on more than one factor. Examination of cross-loadings suggested that more generic concepts related to student/teacher communications (eg. clarity in directions, explanations and grading) tended to conceptually align with more than one SATL assessment domain. For the most part, the patterning of loadings confirmed the original logical classification of SATL indicators by Assessment Domains and represented the best statistical and conceptual fit with the original structure of the SATL.

Descriptive statistics for items and factored dimensions (Domains/Subscales) indicated that instructors received relatively consistent scores (80.7% to 86.4% of the maximum possible score) across SATL Assessment Domains (Table 3). The lowest percentages were noted for the Assessment Domain, Enhancement of Learning, which includes newer assessment indicators targeting the extent to which teaching methods and learning tasks actively involve students in the development of higher order thinking skills. Interestingly and importantly , assessment scores for the various SATL assessment indicators were lower for indicators targeting the teaching and learning of higher order thinking skills than for other assessment indicators. When viewed from a mastery perspective, for example, the percentage of the maximum possible score for all of the SATL thinking skills indicators considered collectively was only 17 percent.

Cronbach Alpha reliability coefficients for each SATL Assessment Domain were as follows: I) Preparation and Classroom Management =.89; 2) Interpersonal Skills =.91; 3) Enhancement of Learning = .96; and 4) Student Evaluation Practices =.93.

Table 4 summarizes Pearson product moment correlation coefficients among the various SATL subscales (Assessment Domains) and students' summary judgements of types of learning emphasized in the classroom and overall class/course evaluation for the total sample (n=339) of students. Intercorrelations among the various SATL subscales ranged from .76 ($p<.0001$) (Enhancement of Learning and Interpersonal Skills) to .20 ($p<.0001$) (Student Evaluation Practices and Overall Evaluation). Intercorrelations between the SATL Overall Course Evaluation component and the four SATL Assessment Domains (Subscales) ranged from .20 ($p<.0001$) to .29 ($p<.0001$). Intercorrelations were also computed between students' perceptions of expected grades at the beginning and conclusion of the course. Interestingly, these were

typically negative in direction, relatively moderate in magnitude, and not statistically significant.

Phase IV: Piloting of the Comprehensive, Collaborative/Feedback Consultation Model

During the 1992-1993 academic year, the SATL student and a parallel faculty-self perceptions measures were further field tested with a larger sample of evening school classes. Responses were received from approximately 1,000 students and 80 faculty from a cross section of courses, again representing a wide variety of educational contexts. This larger SATL data set was subsequently used to further examine construct validity and reliability characteristics. Results of subsequent factor analytic work confirmed initial classification of items by subscales as reported using data in the initial pilot of the SATL.

In addition, participating faculty were sent assessment feedback profiles (Appendix A) which included percentages of students' responses for each point on the rating scale for each item on the SATL. A selected sample of faculty were also sent Course Evaluation Reports (Appendix B) which included a series of histograms depicting various comparisons of assessment data as follows: 1) Comparisons of total faculty ratings with total student ratings; 2) Comparisons of individual faculty ratings with average faculty ratings; 3) Comparisons of student ratings for individual classes with average student ratings across classes.

In addition to this comprehensive feedback, twelve classrooms were assessed by external observers with the STAR and additional piloting of the consultation model was implemented. The sample for this extended field test activity included volunteer faculty (n=7) and graduate teaching assistants and/or instructors (n=5) from a variety of content areas and settings (e.g., nutrition, chemistry lab settings, English, religion, philosophy, etc.). Each of these twelve participants were observed in the classroom and assessments of teaching and learning were made

by trained observers using the STAR. Each also participated in an assessment feedback consultation conference in which the observer assumed the role of change facilitator and the participant typically seemed to assume the role of content expert. Results were somewhat mixed as receptivity to feedback was varied and most participants seemed to be largely concerned with presentation of content rather than with the interaction of elements of content and pedagogy which have implications for student understanding and learning through engagement in higher order thinking processes. Results provided a basis for further development of a consultation feedback model which includes encouragement and engagement of participants in conversations about teaching and learning, collaborative reflection, the use of multiple data sources rather student perceptions data alone, and, finally, emphasis on teaching as it enhances learning for students rather than simply on delivery of content.

DISCUSSION

This study describes four phases of R&D activity occurring over a period of three years to develop and pilot a comprehensive, learning-based assessment and consultation system in higher education settings. Important, and central to this system is the emphasis placed on classroom activities and events essential to the enhancement of teaching and learning. The results of this study are important from four major perspectives. First, they provide empirical support for the construct validity and reliability of a new, comprehensive measure of students' perceptions of characteristics of teaching and learning environments in higher education settings, particularly in classes composed of non-traditional adult learners (typically part time and older students). Results of the factor analyses completed documented reasonable alignment of the SATL assessment indicators with the Assessment Domains in which they had been originally logically

content classified. Initial reliability (internal consistency) estimates for the various Assessment Domains are encouraging, though future investigations of stability characteristics of the SATL should be (and are presently being) explored.

The focus of the SATL on characteristics of teaching and the enhancement of learning makes it rather unique in the higher education learning environment assessment literature. While other student perceptions measures of learning environment characteristics have been developed for use in higher education settings (e.g., Fraser, Treagust, & Dennis, 1986), these do not have as a central focus students' views of the extent to which a set of reasonably research-based indicators of teaching and learning activities and learning environment variables enhance student learning. The SATL is thought to be unique in this regard and in its focus on newer, important assessment concerns such as the teaching and learning of higher order thinking skills as well. From both the formative and summative evaluation of teaching perspectives, the SATL provides a rich alternative to the wide variety of student course evaluation forms currently being used in higher education settings that typically have a simpler and more narrow rating focus on instructor or course characteristics alone. In this regard, the SATL represents a logical extension of newer ideas derived from recent developments of classroom-based assessments of teaching and learning for public school teachers (Ellett, 1990).

Secondly, the results suggest that students' perceptions of the teaching and learning environment as measured by the SATL have little or no relationship to students' beliefs about expected course grades at either the beginning or the end (after final examinations) of a course. This finding suggests that the SATL response format is relatively independent of such student grade expectations and calls into question past criticisms of the use of student perception

measures to assess (evaluate) higher education classrooms. Thus, such student perceptions may be relatively free of the influence of halo and pitchfork effects associated with grades that are commonly believed to influence students' perceptions of the quality of teaching in higher education classrooms. In addition, students' perceptions of the effectiveness of various SATL assessment indicators in these classes showed only very modest, positive relationships to their global, summary course evaluation judgments (r=.20 to .29). This finding seems to lend additional support to the relative independence of the SATL indicator assessments from pitchfork and halo effects in evaluation.

Third, the results showed that the SATL assessment indicators targeting the teaching, development and learning of higher order thinking skills received the lowest scores of all SATL assessment indicators. Indeed, scores for these indicators were somewhat lower than similarly noted low scores obtained through direct classroom-based observations of teaching and learning recently completed in some 6,000 public school classrooms with the STAR (Ellett, Loup, & Chauvin, 1990) throughout Louisiana (Hill, Claudet, Loup, Chauvin, & Naik, 1991). These findings suggest that these higher education learning environments were not characterized by the teaching and learning of higher order cognitive skills. Instead, they were perceived by students as primarily focused on the acquisition of factual knowledge using traditional, receptive teaching and learning modes. To the extent that this finding is generalizable to other higher education institutions and settings, rather serious issues and concerns about the quality of teaching and learning in higher education are raised. Among these is the extent to which faculty in such settings may need to learn more effective teaching methodologies designed to develop higher order thinking among students.

Fourth, and from a more personal and clinical perspective, implementation of a collaborative feedback/consultation model in this study provided considerable insights about an important set of variables that need to be considered in efforts to improve teaching and learning in higher education settings. Results of conferences with faculty and graduate teaching assistants served to identify a variety of metaphors that seem to pervade perceptions of teaching and learning in higher education classrooms. For example, many faculty viewed themselves as essentially "content deliverers" and viewed teaching and learning as essentially moving from "me to thee." Thus, the concerns reflected in feedback consultations typically focused on how faculty might improve their teaching techniques...rather than student learning. Faculty perceptions of the teaching role as "facilitator" or "enhancer" of learning were most evident in conversations with one faculty in a helping professions setting (nurse).

Working with faculty to change these habituated perspectives about learning in higher education settings raises an important set of concerns. A key to such changes may be findings means to increase the legitimacy and perceived importance of information derived from both student evaluation data and direct observation data. In the study reported here, faculty members varied considerably in the extent to which they believed in and valued evaluation results. Some became very concerned about evident low scores (particularly those from the SATL), while others became somewhat defensive about sharing results and appeared very hesitant to discuss alternative strategies and techniques targeting improvements. In this regard, graduate teaching assistants seemed more interested in, and valued the evaluation results more as a means to improve, than did regular faculty members.

Another concern arising from feedback conferences with faculty relates to self reflections

about teaching and learning. These faculty seemed to give priority to a personal focus on teaching...and gave little consideration to either improving student learning...or to establishing better linkages between the quality of teaching and student learning. They seemed to view and define their effectiveness as teachers in isolation of concerns for student learning. The qualitative findings reported here strongly suggest that faculty in higher education settings might well profit from inservice education that focuses on linkages between effective teaching **and** student learning. The SATL data in this study, for example, strongly suggested that students did not believe that faculty and graduate teaching assistants had any knowledge of student levels of misunderstanding of content and subject matter. Developing such understandings, of course, may directly depend upon the ability to use teaching strategies that target informal assessments of student learning <u>during</u> classroom teaching and learning processes.

The assessment system and consultation model provides alternatives to more traditionally used student perceptions instruments that focus on <u>course and professor evaluation</u> as opposed to actual <u>teaching effectiveness as it relates to student learning</u>. The complete assessment system includes measures that triangulate the perspectives of students, external observers, and faculty self-assessments. The SATL instrument used to collect student perceptions data is grounded in and derived from a comprehensive, classroom-based observation system (STAR) originally designed to assess the quality of teaching and learning in classrooms of public school teachers for the purpose of professional certification and improvement (Ellett, Loup, & Chauvin, 1990). Interestingly, the results from the extended field test in Phases III and IV of the study suggest that college/university classrooms share some common strengths and weaknesses pertaining to effective teaching and learning reflected in statewide assessments of public school classrooms

(Hill, Claudet, Loup, Chauvin, & Naik, 1991).

In addition, preliminary findings from piloting of the consultation feedback model with twelve university faculty and assistants and earlier case studies with graduate assistants suggest that triangulated data sources in the assessment system are a viable means of providing meaningful feedback to college/university faculty for improving teaching and learning in higher education classrooms.

The findings from this three-year research and development project to improve teaching and learning in higher education settings shed some light on the traditional rhetoric associated with the difficulty in developing and conducting valid and reliable assessments in higher education classrooms. In addition, the response of faculty participants in this study to this collaborative consultation model suggests that the traditional rhetoric surrounding the importance of matching an observer/assessor's course/content knowledge with that of the faculty member is somewhat suspect. Both the SATL and the STAR appear to assess a variety of important elements associated with effective teaching and learning environments (process variables) that are generic across different content areas in higher education settings.

Finally, the results have implications for using the SATL in the future study of learning environment characteristics and for the development of systematic lines of inquiry relating teaching and learning environment characteristics to student outcomes. Written comments solicited from students and instructors in this study indicated that the SATL is easy to administer, clear in content and focus, and adequately addresses important aspects of a variety of course and instructor variables defining multiple classroom learning contexts in higher education. The SATL seems efficient as a data collection device since it requires approximately 20 minutes to

administer. Additionally, with its response focus on the extent to which elements of teaching and classroom activities enhance learning, rather than on broader notions pertaining to psychosocial characteristics of the learning environment, the SATL might be used as a proxy student perceptions measure of student learning and achievement in future research on learning environment characteristics in higher education settings.

REFERENCES

Aleamoni, Lawrence. (1987). Student rating myths versus research facts. Journal of Personnel Evaluation in Education, 1(1), 111-119.

Berliner, D. (1986, September/August). In pursuit of the expert pedagogue. Educational Researcher, 5-13.

Berliner, D. (1988). Implications of studies of expertise in pedagogy for teacher education and evaluation. New Directions for Teacher Assessment. Proceedings of the 1988 ETS Invitational Conference, Princeton, New Jersey: Educational Testing Service, 39-65.

Brinko, K. (1991). The interactions of teaching improvement. In M. Theall and J. Franklin (Eds.), New directions for teaching and learning: Effective practices for improving teaching. San Francisco, CA: Jossey/Bass, 48, 39-49.

Brophy, J. (1986). Teacher influences on student achievement. American Psychologist, October, 1069-1077.

Ellett, C.D. (1990). A new generation of classroom-based assessments of teaching and learning: Concepts, issues and controversies from pilots of the Louisiana STAR. Working paper, Baton Rouge, Louisiana: Louisiana Teaching Internship and Statewide Teacher Evaluation Projects, College of Education, Louisiana State University.

Ellett, C.D., Garland, J. & Logan, C. (1987). Content classification, synthesis and verification of eight large-scale teacher performance assessment instruments. Project report, Baton Rouge, Louisiana: Teaching Internship Project, College of Education, Louisiana State University.

Ellett, C.D., Loup, K.S., & Chauvin, S.W. (1991a). Development, validity and reliability of a new generation of assessments of effective teaching and learning: Future directions for the study of learning environments. Journal of Classroom Interaction, 26(2), 25-39.

Ellett, C.D., Loup, K. & Chauvin, S.W. (1991b). System for Teaching and learning Assessment and Review (STAR). Assessment manual, Baton Rouge, Louisiana: Louisiana Teaching Internship and Statewide Teacher Evaluation Projects, College of Education, Louisiana State University.

Evans, R.L., & Ellett, C.D. (1990). An initial application of the System for Teaching and learning Assessment and Review (STAR) to the assessment of teaching and learning in selected undergraduate classrooms at Louisiana State University. Technical Report Number 2, Center for Faculty Development. Baton Rouge, Louisiana: Louisiana State University.

Evans, R.L., & Ellett, C.D. (1991). An application of the System for Teaching and learning Assessment and Review (STAR) to the assessment of teaching and learning in selected undergraduate classrooms at Louisiana State University. Technical Report Number 1, Center for Faculty Development. Baton Rouge, Louisiana: Louisiana State University.

Evans, R.L, & Ellett. C.D. (1992). Student Assessment of Teaching and Learning (SATL). Baton Rouge, Louisiana: Division of Continuing Education, Louisiana State University.

Fraser, B., Anderson, G.J. & Walberg, H.J. (1982). Assessment of Learning Environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI) (3rd. Edition), Perth, Australia: Western Australian Institute of Technology.

Fraser, B.J. and Fisher, D.L. (1983). A comparison of actual and preferred classroom environment as perceived by science teachers and students. Journal of Research in Science Teaching, 20, 55-61.

Fraser, B.J. Treagust, D.F. & Dennis, N.C. (1986). Development of an instrument for assessing classroom psychosocial environment in universities and colleges. Studies in Higher Education, II, 43-54.

Fraser, B.J., Treagust, D.F., Williamson, J.C. & Tobin, K.G. (1987). Validation and application of College and University Classroom Environment Inventory (CUCEI). In B. Fraser (Ed.) The Study of Learning Environments: Volume II. Research monograph of the Study of Learning Environments SIG of the American Educational Research Association, Launceston, Australia: University of Tasmania.

Fraser, B.J., & Walberg, H.J. (1991). Educational environments: Evaluation, antecedents and consequences. London: Pergamon Press.

Glaser, R. (1985). Learning and instruction: A letter for a time capsule. In S.F. Chipman, J.W. Segal, & R. Glaser (Eds.), Thinking and learning skills (Vol. 2). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Hill, F.H., Claudet, J.G., Loup, K.S., Chauvin, S.W., & Naik, N.S. (1991). Classroom assessments of 6000 teachers: What do the results show about the effectiveness of teaching and learning?. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois.

Kauchak, D.P., & Eggen, P.D. (1989). Learning and teaching: Research-based methods. Boston: Allyn & Bacon.

Logan, C. & Ellett, C.D. (1988). Development, validity and reliability of a faculty evaluation instrument to measure generic teaching skills and perceived enhancement of student learning. Journal of Personnel Evaluation in Education, 2(1), 65-82.

Loup, K., Ellett, C.D. & Chauvin, S.W. (1992). The System for Teaching and learning Assessment and Review (STAR): A holistic classroom observation alternative to measures of student perceptions for research on classroom learning environments. In D. Fisher (Ed.) The Study of Learning Environments: Volume VI. Research monograph of the Study of Learning Environments SIG of the American Educational Research Association, Launceston, Australia: University of Tasmania, 76-92.

Marcelo, C. (1990). Research on psychosocial environment evaluation at university classrooms: Adaptation of C.U.C.E.I. to the spanish educational context. In H. Waxman & C. Ellett (Eds). The Study of Learning Environments: Volume IV. Research monograph of the Study of Learning Environments SIG of the American Educational Research Association, Launceston, Australia: University of Tasmania.

Marsh. H.W., & Dunkin, M. (1992). Student evaluations of university teaching: A mulit-dimensional perspective. Higher Education: Handbook on Theory and Research, Vol.8. New York: Agathon.

Marsh, H., & Roche, L. (1993). The use of students' evaluations in an individually structured intervention to enhance university teaching effectiveness. American Educational Research Journal. 30(1), 217-251.

Menges, Robert. (1992). Control, autonomy, and the college teacher. Invited address at the American Educational Research Association Faculty Evaluation and Development SIG, San Francisco, CA.

Menges, R., & Brinko, K. (1986). Effects of student evaluation feedback: A meta-analysis of higher education research. ERIC Document Reproduction Service (270-408). Evanston, Illinois: Northwestern University.

Moos, R.H. (1974). The social climate scales: An overview. Palo Alto, CA: Consulting Psychologists Press.

Porter, A.C. and Brophy, J. (1988). Synthesis of research on good teaching: Insights from the work of the institute for research on teaching. Educational Leadership, 74-85.

Purtle, V. (1982). Guidelines for student evaluation of teaching. (Unpublished Manuscript), Baton Rouge, Louisiana: College of Arts and Sciences, Louisiana State University.

Shulman, L.S. (1986, February). Those who understand: Knowledge growth in teaching. Educational Researcher, 15, (2), 4-14.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57, (1), 1-22.

Table 1

Summary of SATL Indicators of Effective Teaching and Learning

Factor Pattern Loadings for a One-Factor Solution

for Each Assessment Domain (n=339)

INDICATOR NUMBER		FACTOR PATTERN LOADING
ASSESSMENT DOMAIN I: Preparation and Classroom Management		
1.	Objectives for the course are clearly communicated.	.63
2.	Objectives for each class are clearly communicated.	.65
3.	Student responsibilities and expectations are clearly explained.	.62
4.	The professor is well-prepared for class.	.58
5.	Class starts and ends on time.	.38
6.	Time during class is efficiently used for teaching and learning.	.60
7.	The course activities over the semester are well-organized.	.64
8.	Class activities help to achieve the objectives of the course.	.68
9.	Assigned readings (textbook and/or other) are meaningful.	.57
10.	Outside assignments help to achieve the course objectives and are integrated with class activities.	.57
11.	Teaching and learning techniques motivate students to learn.	.79
ASSESSMENT DOMAIN II: Interpersonal Skills		
12.	The professor is enthusiastic about teaching this course.	.59

Table 1 (Cont.)

13.	The professor is enthusiastic about students' learning in this course.	.67
14.	The professor maintains a climate conducive to learning.	.69
15.	The professor relates the subject to everyday life.	.65
16.	Students are encouraged to express their own ideas.	.59
17.	Students are encouraged to participate in discussions.	.52
18.	A climate of mutual courtesy and respect is maintained.	.69
19.	The professor demonstrates interest in the progress of individual students.	.71
20.	The professor is willing to provide outside help and guidance.	.65
21.	The professor is sensitive to the needs and feelings of students.	.70
ASSESSMENT DOMAIN III: Enhancement of Learning		
22.	The professor arouses and maintains students' interest in the subject.	.78
23.	Teaching methods stimulate interest in the course.	.83
24.	The professor's speech is audible and easily understood.	.63
25.	Directions and explanations related to course content are clear.	.70
26.	Thought-provoking questions are asked.	.67
27.	Questions are asked that allow students to compare and contrast ideas.	.63
28.	The professor draws students into discussions among themselves.	.59
29.	During classes students are encouraged to interact and learn from one another.	.59

Table 1 (Cont.)

30.	Topics are summarized before moving on to new topics.	.66
31.	Important topics or ideas are summarized at the end of class.	.67
32.	The professor helps students to organize information and understand relationships among the topics in the course.	.72
33.	Difficult material is clearly explained.	.73
34.	Students are encouraged to ask questions in class.	.60
35.	Students' questions are clearly answered.	.72
36.	Learning activities are implemented at an appropriate pace.	.68
37.	Course content is at an appropriate level of difficulty.	.69
38.	Difficult parts of the lesson are emphasized to help students learn.	.79
39.	The professor seems to know when the students do not understand.	.73
40.	When students are confused, the professor clarifies as needed.	.75
41.	Students receive feedback about their learning during the lesson.	.77
42.	Adjustments are made in the lesson, as needed, to help students learn.	.78
43.	Teaching aids (such as visuals) are used in a way that enhances learning and broadens understanding.	.65
44.	Learning materials (such as books, handout, lab equipment) are used in a way that enhances learning and broadens understanding.	.62
45.	Students are encouraged to apply course content to solve problems or understand real life situations.	.60

Table 1 (Cont.)

ASSESSMENT DOMAIN IV: Student Evaluation Practices		
46.	The basis for grading is made clear.	.62
47.	The basis for the course grade reflects an appropriate balance among tests, papers, assignments, participation, etc.	.65
48.	Tests reflect the course content.	.68
49.	The test questions are clearly phrased.	.69
50.	The tests are of reasonable length.	.69
51.	The test questions provide a fair chance for students to demonstrate their knowledge of the subject.	.69
52.	Course assignments are graded fairly.	.69
53.	Tests and assignments are returned in a reasonable length of time.	.64
54.	Sufficient feedback is provided on all graded work.	.70
55.	Students have opportunities to determine their own progress in the course.	.67

Eigen Value = 24.45

Proportion of Total Variance Explained by 1-Factor Solution = 44.45%

TABLE 2

Summary of SATL Indicator Factor Pattern Loadings for a Four-Factor Orthogonal Solution for Each Assessment Domain (n=339)

	FACTORS			
	I	II	III	IV
ASSESSMENT DOMAIN I: Preparation and Classroom Management				
1.				.54
2.				.58
3.				.61
4.				.60
5.				.46
6.				.59
7.				.69
8.	.49			.43
9.	.36			.46
10.	.51			
11.	.52			.51
ASSESSMENT DOMAIN II: Interpersonal Skills				
12.		.57		
13.		.68		
14.		.47		.48
15.		.51		
16.		.77		
17.		.73		
18.		.60		

Table 2 (Cont.)

19.		.60	
20.		.57	
21.		.64	
ASSESSMENT DOMAIN III: Enhancement of Learning			
22.	.47	.41	.39
23.	.53	.38	.42
24.			.53
25.			
26.	.52		
27.	.62		
28.	.70		
29.	.69		
30.	.52		
31.	.60		
32.	.57		
33.	.49	.36	
34.		.59	
35.	.37	.51	
36.	.69		
37.	.55		
38.	.57		
39.	.42	.48	
40.	.39	.46	.36
41.	.63		
42.	.59		
43.	.61		

Table 2 (Cont.)

44.	.58			
45.	.48			
ASSESSMENT DOMAIN IV: STUDENT EVALUATION PRACTICES				
46.			.59	.47
47.			.58	
48.			.69	
49.			.66	
50.			.66	
51.			.69	
52.			.68	
53.			.73	
54.			.71	
55.			.66	
Proportion of Total Variance in Data Set Explained by Each Factor in the 4-Factor Solution	16.8	14.10	13.0	12.1

Proportion of the Total Variance in the Data Set Explained by a 4-Factor Solution = 56.23%

TABLE 3

Summary of Descriptive Statistics for Each Subscale of the SATL Four-Factor Solution for the Total Sample of Students

Subscale	N	Mean	SD	Max. Poss. Score	%M[a] Max. Poss.
Planning and Classroom Management (11)[b]	448	27.8	5.49	33	84.2
Interpersonal Skills (10)	445	25.9	5.07	30	86.4
Enhancement of Learning (24)	445	58.1	13.40	72	80.7
Student Evaluation Practices (10)	435	25.2	5.65	30	83.9

[a] M% Max = Subscale Mean Score/Max. Possible Score
[b] Number of Items in Subscale

TABLE 4

Summary of Intercorrelations Among SATL Subscales for the Total Sample of Students

SUBSCALES	PCM	IS	EL	SEP	TL	OE
Preparation and Classroom Management (PCM)	1.0	.75	.75	.68	.57	.27
Interpersonal Skills (IS)		1.0	.76	.62	.57	.29
Enhancement of Learning (EL)			1.0	.70	.65	.28
Student Evaluation Practices (SEP)				1.0	.46	.20
Types of Learning (TL)					1.0	.23
Overall Evaluation (OE)						1.0

*p<.001 (all)

APPENDIX A

Course Evaluation Report Profiles

COURSE EVALUATION REPORT
COURSE ID NUMBER - 999

PART I
3 = Highly effective in enhancing learning
2 = Effective in enhancing learning
1 = Ineffective in enhancing learning

PREPARATION AND CLASSROOM MANAGEMENT

	%-1	%-2	%-3	Not ans.	No. resp.
1. Objectives for the course are clearly communicated.	3	29	68	26	432
2. Objectives for each class are clearly communicated.	4	29	67	30	428
3. Student responsibilities and expectations are clearly explained.	4	26	70	25	433
4. The professor is well-prepared for class.	2	19	79	27	431
5. Class starts and ends on time.	3	21	75	29	429
6. Time during class is efficiently used for teaching and learning.	6	24	71	27	431
7. The course activities over the semester are well-organized.	4	24	72	31	427
8. Class activities help to achieve the objectives of the course.	5	27	68	29	429
9. Assigned readings (textbook and/or other) are meaningful.	9	37	54	24	434
10. Outside assignments help to achieve the course objectives and are integrated with class activities.	9	34	57	33	425
11. Teaching and learning techniques motivate students to learn.	5	30	65	29	429

INTERPERSONAL SKILLS

	%-1	%-2	%-3	Not ans.	No. resp.
12. The professor is enthusiastic about teaching this course.	2	16	82	28	430
13. The professor is enthusiastic about students' learning in this course.	5	19	77	30	428
14. The professor maintains a climate conducive to learning.	4	21	75	30	428
15. The professor relates the subject to everyday life.	4	21	75	28	430
16. Students are encouraged to express their own ideas.	5	19	76	32	426
17. Students are encouraged to participate in discussions.	4	17	80	30	428
18. A climate of mutual courtesy and respect is maintained.	4	17	80	33	425
19. The professor demonstrates interest in the progress of individual students.	6	26	68	33	425
20. The professor is willing to provide outside help and guidance.	3	21	76	30	428
21. The professor is sensitive to the needs and feelings of students.	5	25	70	32	426

ENHANCEMENT OF LEARNING

	%-1	%-2	%-3	Not ans.	No. resp.
22. The professor arouses and maintains students' interest in the subject.	5	33	63	24	434
23. Teaching methods stimulate interest in the course.	5	33	63	31	427
24. The professor's speech is audible and easily understood.	4	16	80	33	425
25. Directions and explanations related to course content are clear.	5	24	71	28	430
26. Thought-provoking questions are asked.	5	30	65	29	429
27. Questions are asked that allow students to compare and contrast ideas.	5	33	61	32	426

BEST COPY AVAILABLE

PART I
3 = Highly effective in enhancing learning
2 = Effective in enhancing learning
1 = Ineffective in enhancing learning

ENHANCEMENT OF LEARNING

	%-1	%-2	%-3	Not ans.	No. resp.
28. The professor draws students into discussions among themselves.	11	31	58	33	425
29. During classes students are encouraged to interact and learn from one another.	14	30	56	34	424
30. Topics are summarized before moving on to new topics.	7	37	56	30	428
31. Important topics or ideas are summarized at the end of class.	8	42	50	30	428
32. The professor helps students to organize information and under-stand relationships among the topics in the course.	5	32	63	30	428
33. Difficult material is clearly explained.	6	31	63	29	429
34. Students are encouraged to ask questions in class.	3	23	74	30	428
35. Students' questions are clearly answered.	5	27	69	32	426
36. Learning activities are implemented at an appropriate pace.	4	38	58	34	424
37. Course content is at an appropriate level of difficulty.	5	33	62	31	427
38. Difficult parts of the lesson are emphasized to help students learn.	8	35	58	32	426
39. The professor seems to know when the students do not understand.	5	32	63	29	429
40. When students are confused, the professor clarifies as needed.	3	30	68	28	430
41. Students receive feedback about their learning during the lesson.	7	37	56	32	426
42. Adjustments are made in the lesson, as needed, to help students learn.	7	33	59	31	427
43. Teaching aids (such as visuals) are used in a way that enhances learning and broadens understanding.	13	33	54	37	421
44. Learning materials (such as books, handout, lab equipment) are used in a way that enhances learning and broadens understanding.	6	31	63	34	424
45. Students are encouraged to apply course content to solve problems or understand real life situations.	6	32	62	34	424

STUDENT EVALUATION PRACTICES

	%-1	%-2	%-3	Not ans.	No. resp.
46. The basis for grading is made clear.	6	23	70	28	430
47. The basis for the course grade reflects an appropriate balance among tests, papers, assignments, participation, etc.	6	26	68	29	429
48. Tests reflect the course content.	6	25	69	34	424
49. The test questions are clearly phrased.	11	27	62	43	415
50. The tests are of reasonable length.	5	32	63	39	419
51. The test questions provide a fair chance for students to demonstrate their knowledge of the subject.	10	26	64	40	418
52. Course assignments are graded fairly.	5	24	72	39	419
53. Tests and assignments are returned in a reasonable length of time.	3	17	81	32	426
54. Sufficient feedback is provided on all graded work.	7	25	68	29	429
55. Students have opportunities to determine their own progress in the course.	5	34	61	33	425

BEST COPY AVAILABLE

PART II - TYPES OF LEARNING

4 = Very much emphasis
3 = Much emphasis
2 = Some emphasis
1 = No emphasis

	%-1	%-2	%-3	%-4	Not ans.	No. resp.
56. learning factual information	5	20	34	41	28	430
57. developing concepts	3	14	37	47	31	427
58. understanding and applying principles and rules	5	15	33	46	33	425
59. understanding and applying theories	7	17	35	41	34	424
60. critical analysis and/or problem solving	8	16	35	42	31	427
61. creative thinking	12	15	28	46	34	424
62. developing knowledge of self and others	14	18	29	39	34	424
63. developing professional, career, and job-related skills	15	21	29	35	30	428
64. developing written communication skills	22	18	25	35	37	421
65. developing oral communication skills	24	23	22	31	33	425

PART III - OVERALL EVALUATION

	average	Not ans.	No. resp.
66. How would you grade the quality of teaching in this course?	-71.78	54	404
67. What was the contribution of the course to your learning?	-74.77	52	406
68. How would you grade this course overall?	-73.21	54	404

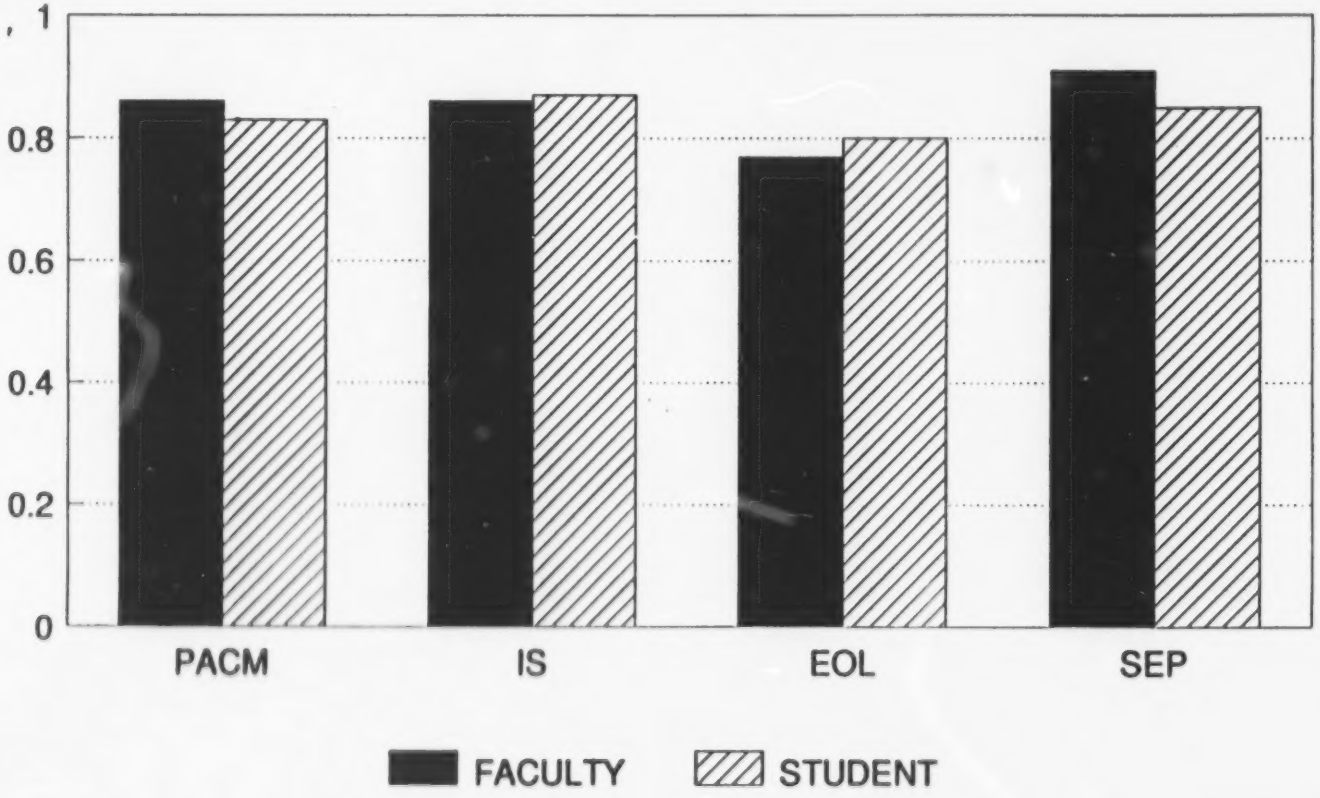
BEST COPY AVAILABLE

APPENDIX B

Course Evaluation Report Histograms

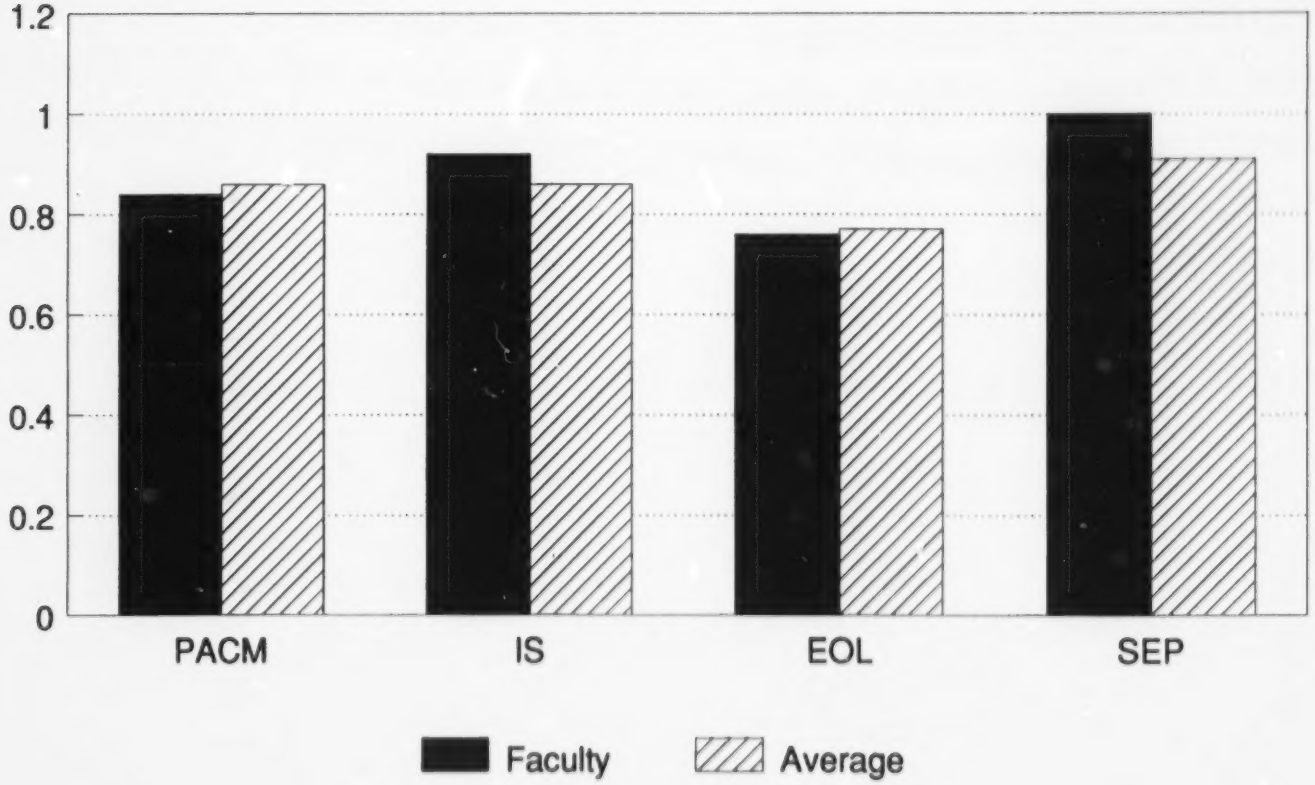
Course Evaluation Report

Total Faculty-Student



Course Evaluation Report

Individual Faculty



Course Evaluation Report

Student (Class ID:)